



Munich Personal RePEc Archive

Bayesian posterior prediction and meta-analysis: an application to the value of travel time savings.

Moral-Benito, Enrique

31. December 2008

Online at <http://mpra.ub.uni-muenchen.de/12861/>
MPRA Paper No. 12861, posted 19. January 2009 / 23:29

BAYESIAN POSTERIOR PREDICTION AND META-ANALYSIS: AN APPLICATION TO THE VALUE OF TRAVEL TIME SAVINGS

ENRIQUE MORAL-BENITO*
CEMFI

DECEMBER 2008

Abstract

In the evaluation of transportation infrastructure projects, some non-tradable goods such as time are usually key determinants of the result. However, obtaining monetary values for these goods is not always easy. This paper introduces a novel approach based on the combination of bayesian posterior prediction and meta-analysis. This methodology will allow to obtain predictive distributions of the monetary values for this type of goods. Therefore, uncertainty is formally considered in the analysis. Moreover, the proposed method is easy to apply and inexpensive both in terms of time and money. Finally, an application to the value of travel time savings is also presented.

KEYWORDS: Bayesian Prediction, Meta-Analysis, Uncertainty,
Value of Travel Time Savings.

1 INTRODUCTION

The Value of Travel Time refers to the cost of time spent on transport, including waiting as well as the actual travel. It includes costs to consumers of personal (unpaid) time spent on travel, and costs to businesses of paid employee time spent in travel. Travel time is one of the largest categories of transport costs, and time savings are often the greatest benefit of transport projects such as new and expanded roadways, and public transit improvements. The Value of Travel Time

*Contact: CEMFI, c/ Casado del Alisal 5, 28014, Madrid, Spain. Phone: +34 914 290 551. Fax: +34 914 291 056. e-mail: enrique.moral@gmail.com, emoral@cemfi.es

Savings -henceforth VTTS- refers to the benefits from reduced travel time. As previously stated, in most of the transportation projects travel time savings represent the most important source of social benefits. Therefore, VTTS is considered a key determinant of the evaluation results.

On the other hand, there are other important sources of user benefits that are based on non-tradable goods. The value of a statistical life -henceforth VSL- is the best example in this category. Public choices about safety in a democratic society require estimates of the willingness of people to trade off wealth for a reduction in the probability of death. Estimates of these trade-offs are used in evaluating environmental issues, public safety in travel, medical interventions and in many other areas. It has become common to call this trade-off the value of a statistical life (VSL). VSL is therefore a measure that is widely used for the evaluation of transportation infrastructure projects that, among other things, reduce the risk of fatal accidents. The literature on the estimation of VSL accross the world is huge, for instance Viscusi and Aldy (2003) provide a good survey.

In order to determine the VTTS (or the VSL) for a particular project, the best option is to conduct a specific study among the potential users of the "project". However, this alternative is always very expensive both in terms of time and money. Hence, in practice, the usual approach is to determine the VTTS of the project under evaluation according to recommended values at a national level (see for example UNITE conventions) or by simply imputing the values obtained in previous similar studies. In both cases, a single VTTS (or VSL) is employed. As a consequence, in those approaches uncertainty is not taken into consideration. The imputed VTTS (or VSL) is considered as the true value without any doubt. However, in reality, this imputed value is only an estimate, and uncertainty is inherent in any estimation procedure. Therefore, the existent uncertainty when imputing both VTTS and VSL should be considered in the evaluation process.

Against this background, this paper introduces a meta-analytical approach that can be easily applied to any project under evaluation. The method is meta-analytical since it will review previous scientific studies. Meta-analysis represents an appropriate technique in order to study VTTS or VSL, given that it can summarize the results of very heterogeneous studies. Moreover, the approach will allow to obtain the whole distribution of the VTTS (or VSL) for a particular project by means of bayesian methods. By doing so, uncertainty is formally considered in the analysis and more reliable results will be obtained as pointed out by de Rus

(2008) among others.

The remainder of the paper is organized as follows. [Section 2](#) describes the approach. Firstly I explain bayesian posterior prediction, then I briefly summarize the concept of meta-analysis and I introduce how to combine both methods. In [Section 3](#) I present an example of how the methodology can work in practice. In particular, I show how we can obtain the distribution of VTTS for a particular project under evaluation. The final section concludes.

2 METHODOLOGY

Since the approach I present in this paper is a combination of two different techniques, I will proceed to explain both concepts separately. Then, I will present how they are combined.

2.1 BAYESIAN POSTERIOR PREDICTION

Bayesian econometrics is the systematic use of a result from elementary probability, Bayes' theorem. Suppose we have a model given by $f_y(y; \theta)$, where y represents the data and θ the parameters. The object of interest from an econometric perspective is the vector of parameters θ . The logic of bayesian inference is to apply Bayes' theorem such that:

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad (1)$$

where $p(\theta|y)$ is referred to as the *posterior density*, $p(y|\theta)$ is the *likelihood* function of the data given the parameters and $p(\theta)$ is the *prior density* of the parameters. In the present case, like in most econometrics, prediction is a major concern. That is, given the observed data, y , the econometrician may be interested in predicting some unobserved data y^* . In our case, the observed data y will be the different estimates of the VTTS (or VSL) for projects of previous studies and their characteristics. The unobserved data y^* that we want to predict will be the VTTS (or VSL) for new projects under evaluation.

The bayesian reasoning argues that uncertainty about the unobserved elements (y^*) are summarized by a conditional probability statement. That is, prediction should be based on the **posterior predictive density** $p(y^*|y)$ given by:

$$p(y^*|y) = \int p(y^*|y, \theta)p(\theta|y)d\theta \quad (2)$$

The predictions we would obtain by applying bayesian posterior prediction¹ are in a different spirit of those obtained by classical methods. The important difference is that with the bayesian approach, we predict the whole distribution of unobserved data instead of a single data point with some standard error. This is the main advantage of the approach presented in this paper with regard to the treatment of uncertainty.

2.2 META-ANALYSIS

Meta-analysis is defined as the process or technique of synthesizing research results by using various statistical methods to retrieve, select, and combine results from previous separate but related studies. Given that a meta-analytical analysis includes more observations than any particular study included in the analysis, its statistical power is higher.

A discussion of the enormous literature on meta-analysis is outside the scope of this paper. Hunter and Schmidt (2004) is a good reference for those readers interested in meta-analysis *per se*.

The focus of this paper is the meta-analysis of different VTTS (and VSL) estimates. As mentioned in the introduction, the literature on both VTTS and VSL is enormous. Thus, once we have compiled enough data of different VTTS and VSL studies, we will be able to carry out a meta-analysis in order to summarize all the information in a rigorous statistical manner.

2.3 META-ANALYTICAL BAYESIAN POSTERIOR PREDICTION

The combination of the two techniques described above allows to obtain the whole distribution of the monetary values of non-tradable goods such as time. The resultant procedure can be denominated Meta-Analytical Bayesian Posterior Prediction.

The word prediction refers to the fact that we will obtain (or forecast) non-observed values. That is to say, based on a sample of observed VTTS (or VSL) from previous analyses, we will be able to predict the VTTS (or VSL) for new projects under evaluation that had never been analyzed before. The method also uses bayesian methods that will allow us to incorporate uncertainty from the very beginning in a more natural way than classical approaches. Finally, given the

¹For more details about the bayesian methods presented here refer to Lancaster (2004)

compilation of data from previous studies and the use of statistical techniques in order to combine this information, the approach is also called meta-analytical.

I next turn to formally introduce the methodology. We depart from a linear regression model:

$$y_i = x_i' \beta + v_i \quad (3)$$

where $i = 1, \dots, N$ refers to the N previous studies for which we have data. y_i is the value of the non-tradable good (*i.e.* VTTS or VSL) reported in the study i and x_i is the $k \times 1$ vector of observable characteristics of study i . (for example trip purpose or mode in the case of VTTS). By stacking the N observations in vectors we can rewrite (3) in matrix form:

$$Y = X\beta + V \quad (4)$$

where now, Y and X are a $N \times 1$ vector and a $N \times k$ matrix of data respectively. β is a $k \times 1$ vector of parameters and V is the $N \times 1$ vector of disturbance terms. Given the model in (4), we can now turn to the application of bayesian posterior prediction. For this purpose we follow a sequential procedure based on three steps: **(i)** Elicitation of the likelihood function for the data and the prior distribution for the parameters. **(ii)** Given the likelihood and the prior, obtain the posterior distribution of the parameters, and finally, **(iii)** solve the integral in (2) in order to obtain the posterior predictive distribution.

First of all, I propose to assume that the error term in (4) follows a multivariate normal distribution with zero mean and variance-covariance matrix given by $\sigma^2 I_N$:

$$V \sim N(0, \sigma^2 I_N)$$

The previous assumption implies that the likelihood is given by:

$$p(Y|\beta, \sigma^2) = \frac{\sigma^{-\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \left\{ \exp \left[-\frac{1}{2\sigma^2} (Y - X\beta)'(Y - X\beta) \right] \right\} \quad (5)$$

In many situations, when we face the problem of choosing a prior distribution for the parameters of a model, we have very little (if any) prior information for such task. This is also the situation we are facing here, and therefore I propose to assume a non-informative (or diffuse) prior for the parameters. As the likelihood function in (5) belongs to the family of normal-gamma distributions, I will elicit a diffuse prior for the parameters by assuming a normal-gamma distribution with infinite variance. As the variance is a measure of uncertainty, by fixing it to infinity, we are assuming that we do not have any prior information.

Given the likelihood and prior proposed above, and by using Bayes' theorem as in (1), the posterior distribution of the parameters is²:

$$\beta, \sigma^2 | y \sim NG(\overline{\beta}, \overline{\Sigma}, \overline{s}^{-2}, \overline{\nu}) \quad (6)$$

where since we have employed diffuse priors, the overlined parameters of the posterior distribution are given by:

$$\begin{aligned} \overline{\Sigma} &= (X'X)^{-1} \\ \overline{\beta} &= \hat{\beta}_{OLS} = (X'X)^{-1}X'Y \\ \overline{\nu} &= N \\ \overline{s}^2 &= \frac{N-k}{N}s^2 \end{aligned}$$

Moreover:

$$s^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{N - k}$$

We are now ready to obtain the predictive distribution of our interest by solving the integral in (2)³. As shown, for example in Koop (2003), with the likelihood and the prior presented above, this integral can be solved analytically. Moreover, the resultant predictive distribution is a Student's t -distribution defined by the following parameters:

$$y^* | y \sim t(X^{*'}\overline{\beta}, \overline{s}^2 [1 + X^*\overline{\Sigma}X^*], \overline{\nu}) \quad (7)$$

where X^* is the $k \times 1$ vector of characteristics of the project under evaluation. Therefore, by simply compiling some information about our project (the X^* vector) and applying (7), we will easily obtain the predictive distribution for the monetary values of the non-tradable goods we are interested in. In the application we will see the kind of information that we will need depending on the availability and the value of interest.

2.4 TRUNCATION OF THE PREDICTIVE DISTRIBUTION

In the previous subsection we have obtained the predictive distribution for monetary values of non-tradable goods such as time. However, up to this point we have only made use of the rules of mathematics and probability without any economics. For example, the resultant t -distribution has support in all the real line. Thus,

²See Koop (2003) for a proof of this result

³Note that in this case the parameters of the model are $\theta = (\beta, \sigma^2)$

this result would imply that we assign some probability mass to negative values of the travel time savings. This is an awkward property.

From previous studies we know that the VTTS (or the VSL) must be between some limits. For example it can not take a negative value in any case. Therefore, for solving this problem, I propose to work with the truncated version of the predictive distribution in (7). The truncated distribution of a random variable x between two points a and b is defined by:

$$f(x \mid a < x < b) = \begin{cases} \frac{f(x)}{F(b)-F(a)} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

Given the above, once we have obtained the parameters of the predictive t -distribution of interest, we will also need to compute its truncated version between the points (VTTS or VSL) with some economic sense from previous studies, for instance, between the maximum and the minimum of all the values previously estimated.

2.5 EVALUATING PREDICTIVE ABILITY

Prediction occupies a central position in the methodology proposed in this paper; hence, evaluating predictive ability is a fundamental concern. Reviews of the forecast evaluation literature, such as Diebold and Lopez (1996), reveal that most attention has been paid to evaluating point forecast. However there is an insightful evaluation method for density forecast. Based on the probability integral transform introduced by Rosenblatt (1952), I propose here to evaluate our predictive distributions in practice by employing the idea of Diebold et al. (1998).

Consider the following ingredients:

1. A sequence of observations $\{y_i\}_{i=1}^N$ generated by the functions $\{f_i(y_i) \mid \Omega_{-i}\}_{i=1}^N$ where $\Omega_{-i} = y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_N$.
2. A group of predictive densities for these observations $\{p_i(y_i)\}_{i=1}^N$ obtained by using the meta-analytical bayesian posterior prediction proposed in this paper.
3. The probability integral transform of the predictive density $p_i(y_i)$ that is given by:

$$z_i = \int_{-\infty}^{y_i} p_i(u) du = P_i(y_i)$$

The idea of the evaluation method is simple: if the sequence of predictive densities coincides with the sequence of the functions that generated the data, that is to say, if the predictive densities coincide with the real densities, then, the sequence of probability integral transforms must satisfy:

$$\{z_i\}_{i=1}^N \stackrel{\text{iid}}{\sim} U(0, 1) \quad (9)$$

Therefore, we only need to obtain the predictive density for an existent project by using the information in all the other projects and excluding its own information. This means that the matrix of data X that we will employ in the computation of the parameters of the posterior distribution given by (7) will have $N - 1$ rows instead of N and $\bar{\nu}$ will be equal to $N - 1$. If we obtain in this way the predictive distribution for all the N projects in our sample, we can then compute the sequence of N probability integral transforms and test the hypothesis in (9) by means of, for instance, a Kolmogorov-Smirnov test.

3 APPLICATION

3.1 THE FICTIOUS PROJECT

As pointed out by Campos et al. (2007), High Speed Railway (HSR) is currently considered as one of the most important technological breakthroughs in passenger transportation developed in the second half of the 20th century.

Imagine that a High Speed Railway (HSR) Spanish project is under evaluation. In particular, this project would link two cities, Madrid in the center and Oviedo in the north of Spain. Moreover, imagine that after analyzing the cost of building the infrastructure, the operating and maintenance costs and the current demand and its projections, we have now to analyze user's costs. More precisely, we need to assign a monetary value for the travel time savings of all potential users of the HSR between Madrid and Oviedo. If these potential users take the HSR instead of the bus, the journey would be, for instance, one hour shorter. Therefore, this time saving would be an important source of social benefits given by the HSR project and must be analyzed in monetary terms as part of the Cost-Benefit Analysis (CBA). Then, if we are willing to consider uncertainty as part of the process, one question arises: what is the distribution of the monetary VTTS for those users? The meta-analytical bayesian posterior prediction introduced in this paper will allow us to answer this relevant question.

3.2 META-ANALYTICAL DATA

We have data of 90 different studies on VTTS carried out in 15 different countries. In particular we take the VTTS reported in each of the studies together with the characteristics of the study (country, year, trip purpose and mode). This data has been compiled by Zamparini and Reggiani (2007) and more details about its compilation as well as some descriptive statistics can be found in the original source. It is important to note that VTTS values are all measured as a percentage of the hourly wage in order to facilitate comparisons of all of the sampled studies.

In order to complete the dataset and base our analysis in a richer source of information, I have also collected country-level data on GDP in constant prices (2000 USD) and population from Penn World Tables 6.2. Per capita GDP in a country will obviously affect the VTTS of that country. It is reasonable to think that in a rich country such as USA, the opportunity cost of its population is high and then monetary value of its inhabitants' time will be higher than in a poorer country like Lesotho. Descriptive statistics of per capita GDP in thousands of 2000 USD collected for the 15 countries in the sample at the time the studies were conducted are shown in Table 1. As expected, the pattern is the same as in Zamparini and Reggiani (2007). The lowest per capita GDP in the sample corresponds to a study carried out in the United Kingdom in 1959, while the highest per capita GDP refers to a study conducted in Norway in 1997.

Table 1: Descriptive Statistics for per capita GDP.

	Per capita GDP thousands of 2000 USD
MEAN	18.65
MEDIAN	19.03
MAXIMUM	31.38
MINIMUM	10.13
STANDARD DEVIATION	4.52
OBSERVATIONS	90

All of the variables described above are considered as determinants of the VTTS for a given study. Moreover, all of this information seems to be easy to collect for a given project in a given country⁴. Therefore, given the approach,

⁴Note that this variables would comprise the vector X^* in (7) and they would allow us to obtain the distribution of $VTTS^*$

we only need to collect this information in order to obtain the predictive density distribution of the VTTS for a particular project under evaluation.

3.3 THE VTTS PREDICTIVE DISTRIBUTION

Table 2 shows the main HSR project characteristics we need for estimating the VTTS predictive distribution. For illustrative purposes, I will estimate the distribution for those users whose trip purpose is employer's bussiness, but the approach can be easily extended to other trip purposes.

Table 2: Project Characteristics.

COUNTRY	Spain
YEAR	2008
MODE	Train
TRIP PURPOSE	Bussiness
PER CAPITA GDP	29,960 USD

Project characteristics presented in Table 2 comprise the vector X^* in (7). The other parameters that define the predictive t -distribution in (7) are as follows:

$$\begin{aligned}
\text{Mean} &= X^{*\prime} \bar{\beta} = 130.87 \\
\text{Variance} &= \bar{s}^2 (1 + X^* \bar{\Sigma} X^{*\prime}) = 2976.22 \\
\text{Degrees of freedom} &= \bar{\nu} = 90
\end{aligned}$$

Given the above, if we call $VTTS^*$ the value of travel time savings for this particular fictious project, our posterior predictive density obtained with the method proposed in this paper is:

$$VTTS^* \sim t(130.87, 2976.22, 90) \quad \text{if} \quad 13 < VTTS^* < 342 \quad (10)$$

where 13 and 342 are the truncation points given by the minimum and the maximum VTTS of all the previous studies.

Therefore, given the obtained predictive distribution in (10), it is only necessary to simulate the different scenarios in our CBA according to this distribution for the VTTS. A method of simulating data from a truncated distribution is described in the [Appendix A.1](#).

I now turn to evaluate the predictive ability of the method in this particular context. Firstly, I obtain the VTTS predictive distribution for all the 90 studies

in the dataset using the information available in the remaining 89 studies. Then, I compute the probability integral transform of all the 90 densities. In this way, I have a sequence $\{z_i\}_{i=1}^{90}$ and I have to test whether this sequence is independent and identically distributed as a $U(0,1)$ distribution. For this purpose, I employ a Kolmogorov-Smirnov test whose null hypothesis is that the sequence is not independent and identically distributed as a $U(0,1)$ distribution. The obtained p-value is 0.00, hence we can clearly reject the null. Therefore, given this result, we are able to conclude that the predictive ability of the method is satisfactory, at least in this particular application.

For additional insights, we can also have a look to the graph of the predictive distribution:

Figure 1: Truncated Predictive Distribution of VTTS

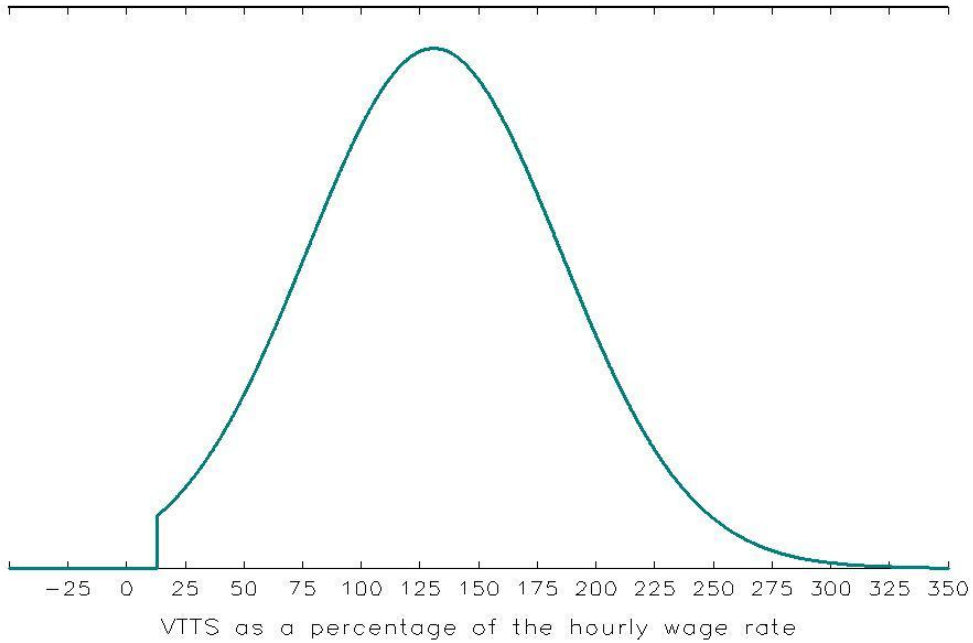


Figure 1 makes clear that uncertainty is present in the analysis. Moreover, given the approach more reliable affirmations can easily be done. For example, in this particular case, we can say that the probability of the VTTS being bigger than the hourly wage rate is 73.5% when the purpose of travel is bussiness in the HSR between Madrid and Oviedo.

4 CONCLUDING REMARKS

In the Cost-Benefit Analysis of different transportation infrastructure projects, some non-tradable goods such as time or lifes, are usually key determinants of the evaluation results.

In spite of being considered in other stages of the CBA, uncertainty is not usually taken into account when imputing values to those non-tradable goods. This is so because conducting specific studies for each project is expensive and the convential approach is to assign recommended single values.

In this paper, a method for estimating the whole distribution of the values of non-tradable goods has been presented. Therefore, it allows to obtain the whole distribution of predicted values taking uncertainty into consideration in a natural manner. The method is appealing because it is easy to apply, requires only little information and is inexpensive in terms of both time and money. Moreover, the predictive ability of the approach can be easily tested.

Finally, for additional insights, an application to the VTTS for a fictious project was presented. The obtained predictive distribution looks reasonable and its predictive capacity is shown to be satisfactory.

Acknowledgements

I thank Ofelia Betancor, Javier Campos, Carlos González-Aguado, Ginés de Rus and Ignacio Sueiro for helpful comments and suggestions. Funding from the CEDEX Ministerio de Fomento through project PT-2007-001-02IAPP is also gratefully acknowledged.

A APPENDIX

A.1 RANDOM DRAWS FROM A TRUNCATED DENSITY

Given a random variable $x \sim f$, I next explain how to extract random draws in practice from its truncated density $f(x|a < x < b)$.

First of all, the truncated cumulative density function (cdf) of x is given by:

$$\frac{F(x) - F(a)}{F(b) - F(a)}$$

We know that the cdf of a random variable is always between 0 and 1 and is uniformly distributed. Therefore, we can generate the random draws from the truncated density as follows:

$$\begin{aligned} u &\sim U[0, 1] \\ x &= F^{-1}[uF(b) + (1 - u)F(a)] \end{aligned}$$

If the variable is not standardized with mean μ and variance σ^2 , we would simulate as follows:

$$x = \mu + \sigma F^{-1}[uF(b) + (1 - u)F(a)]$$

Given the above, in practice we only need to simulate a sequence of numbers from a $U(0, 1)$ distribution, and then apply the previous formulas.

References

- [1] Campos, J., de Rus, G., Barron, I., 2007. Some stylized facts about high speed rail. A review of HSR experiences around the world. Unpublished Manuscript.
- [2] de Rus, G., 2008. Analisis Coste-Beneficio, Ariel Economia.
- [3] Diebold, F., Gunther, T., Tay, A., 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review*, 39, 863-883.
- [4] Diebold, F., Lopez, J., 1996. Forecast evaluation and combination. *Handbook of statistics*, 241-268.
- [5] Hunter, J., Schmidt, F., 2004. Methods of Meta-analysis: Correcting error and bias in research findings, SAGE.
- [6] Koop, G., 2003. Bayesian Econometrics, Wiley-Interscience.
- [7] Lancaster, T., 2004. An Introduction to Modern Bayesian Econometrics, Blackwell Publishing.
- [8] Rosenblatt, M., 1952. Remarks on multivariate transformation. *Annals of mathematical statistics*, 23, 470-472.
- [9] Viscusi, W., Aldy, J., 2003. The value of a statistical life: a critical review of market estimates throughout the world. *NBER Working Papers*, No. 9487.
- [10] Zamparini, L., Reggiani, A., 2007. Meta-Analysis and the value of travel time savings: a transatlantic perspective in passenger transport, *Networks and Spatial Economics*, 7, 377-396.